Survey Solutions
COMPUTER-ASSISTED PERSONAL INTERVIEWING
Michael Lokshin
Benefits of CAPI

- Improve timeliness of data collection
- Ensure data quality and comparability
- Allow collection of new types of information/data
- Cost-effective, sustainable solution for NSOs
Results:

From data collection to policy: 4-5 months
GPS in action

Plots can be easily and precisely measured by simply walking around the area with the tablets.

Ability to measure distances to banks, schools, roads, wells. Maps used for optimal routing.
Sensors

- Water quality testing
- Portable scales
- Vision tests
- Wearable devices
- Multi-purpose sensor
- Soil Fertility sensor
Background: Survey Solutions

- 2011: Comprehensive Assessment of CAPI software products is commissioned to University of Maryland by the WB.
  
  ... no existing software provides exactly the right mix of features necessary for the sort of surveys conducted by World Bank and its clients.


- 01.2015: Surveys and Methods team is formed in DECRG. Complete value chain for survey data collection.

- 02.2017: **Survey Solutions 5.17** is released
CAPI System Requirements

• Simple yet flexible system for the non-expert users. Typical clients – National Stat Offices

• Functionality for
  – **data capturing**: entering data on a tablet
  – **survey management**: managing teams of enumerators
  – **data management**: data aggregation, versioning, reporting

• Tablet-based with ability to display and navigate through multi-level large questionnaires.
• Support of panel surveys and complex validation algorithms.
• Cost effective system that can be used and supported by NSOs without external TA.
Other CAPI systems:

- **CSPro** – free, closed source. resides in Bureau of Census USA. Funded by USAID.
- **ODK** – free, open source software funded by USAID; UWashington, UC Berkeley, UC Davis.
  - SurveyCTO, license fee
- **Blaise** – Statistics Netherlands, license fee.
- **SurveyBe** – closed source, license fee.
- **SurveyToGo** – closed source, license fee.
\[ DQ = \alpha DC \cdot \beta DM \cdot \gamma SM \]

DQ – data quality
DC – data capturing
DM – data management
SM – survey management

Survey Solutions: \( \alpha \approx \beta \approx \gamma \)
CAPI: Expert-centered approach

• Pros:
  – A flexible system that can accommodate specific survey requirements.

• Cons:
  – Very steep learning curves
  – Custom solutions cannot be easily modified for different environments
  – High costs
  – High requirements for the users.
  – Questionable sustainability in the high turnover environments typical for NSO.
• Pros:
  – High levels of standardization, solutions could be easily adapted for different surveys in different countries.
  – Lower requirements for the users, fast learning curve
  – High sustainability
• Cons:
  – Less flexible compared to custom solutions of expert-centered software.
Survey Solutions: hybrid approach

- Sustainable, low-cost system for NSO
- Simple, flexible interface for questionnaire development and testing.
- Tablet interface allows easy navigation through complex questionnaires.
- Standardized survey management protocol based on the best practices of data collection
- Intuitive, informative survey status reporting, survey maps.
- Yet, powerful language for data validation and control of questionnaire flow.
Prices of tablets start from $60: cheaper than paper questionnaires

Questionnaire delivery, Changes or corrections at no cost
Survey cost per household (US$ 2005 PPP)

Legend:
- Red: Implementation
- Gray: Data Processing
- Black: Data analysis and reports
- Blue: Other cost
- Green: Total survey cost

Source: Africa Statistics Team.
Select USD or PPP cost
US$ 2005 PPP

Include Preparation
Yes

Include Implementation
Yes

Include Data
Yes

Include Others
Yes

Include Reporting
No

Legend
- Implementation cost
- Preparation cost
- Data processing cost
- Reporting cost
- Others cost
- Total survey cost

Average at 2,659,384

Source: Africa Statistics Team.
Free software but not free survey

Software might be free but the TA support is very expensive:

Competitor #1: Proposal for a survey in Sri Lanka, 24,000 hhs – $340K
Competitor #2: Actual cost of a survey in Uganda, 10,000 hhs - $350K

Survey Solutions: Cost varies from about $5,000 to 0,000s for large surveys.
Main differences from other systems

- **Out-of-the-box solution for survey data collection**: data capturing, data management, and survey management. No software on the market provides such a package. All other system focus mostly on data capturing.

- **Minimum TA; Focus on Capacity Building**: Survey Solutions is designed to minimize the TA. Lowest learning curve. Other systems require constant and significant TA. Expert versus User centered approach.

- **Designed for large surveys**: Survey Solutions is specifically designed for LSMS and HBS-types of surveys: Nested Rosters, Cascading and Linked questions, Roster-specific validations. Online collaboration.

- **Data security**: Survey Solutions allows storing data on the local servers of NSO thus complying with the local data privacy and anonymity laws. No other software used by the NSOs has such functionality (SurveyCTO can do that for a high fee, but we know of no NSO that is using it – different market (Kenya)).
Survey Solutions: data capturing

- Large questionnaires: several thousand questions
- All standard types of questions:
  - text, numeric, date; multi-choice; dynamic lists;
  - Linked questions: Whose cow is it? User can select from the list of household members.
  - GPS location; Time; Barcode; Binary files (pictures)
  - **HQ/Supervisor-filled questions**
- Rosters: can be generated from:
  - Fixed lists; dynamic lists; numeric; multi-choice questions.
  - **Nested rosters with unlimited degree of nesting**
- Interviewer Comments on a question and interview
- Question Instructions
• **Improve data quality by correcting survey process during the field operation.**
  • System records all events with timestamps that happen on a tablet: Data entry, Data correction, Responsibility changes, etc.
  • Analysis of time per interview, time per question, section.
  • Changes in productivity over time, for different interviewers, teams
  • Quality control, monitoring and evaluation
Paradata: improving data quality
SURVEY PIPELINE: 300+ SURVEYS IN 81 COUNTRIES

- Africa: Malawi (2), Uganda (4), Tanzania, Togo, Benin, Madagascar (3), Niger, Nigeria, Cote D’Ivoire (4), Zambia (3), South Africa, Ghana, South Sudan, Mozambique, Ethiopia, Kenya, Rwanda
- MENA: Djibouti (4), Kuwait, Morocco, Tunisia
- SAR: Bhutan (5), Nepal (3), India (2), Pakistan (2), Bangladesh
- EAP: Thailand (2), Myanmar (15), Tonga, Vanuatu, China, Australia, Malaysia, Philippines, Japan
- ECA: Armenia, Azerbaijan, Kirgizstan, Tajikistan, France
- LAC: St. Lucia (4), Paraguay, Belize, Nicaragua
SURVEY SCOPE

• **Survey types:** LSMS, HBS, LFS, Enterprise Survey, EDU, Health

• **Large Clients:** India NSSO; Stats SA; Thailand NSO; FAO; AfDB; BRAC; OPM; Mathematica

• **Largest survey:** SA – 2,600,000 households, 12,000 enumerators, 500 supervisors, 60K questionnaires per day

• **Questionnaire:** Malawi – 3,000+ questions, 94 rosters

• **Planned surveys:**
  - Indonesia Sakarnas 200,000
  - Ethiopia, Uganda, Malawi LSMS
  - Nigeria LSMS
  - WAEMU Household surveys in 8 countries of West Africa
  - Many others
Real time status of interviews
Map of the survey
Monitor the survey by checking the GPS location of where and when the interview took place.